

*Katz-Lacabe et al v. Oracle America, Inc.*, N.D. Cal. 3:22-cv-04792 (Hon. Richard Seeborg)

**Joint Statement re: Discovery Disputes**

**I. PLAINTIFFS' POSITION**

Pursuant to the Case Management Scheduling Order, Dkt. No. 34 ¶ 3, and after the parties extensively met and conferred regarding Plaintiffs' First Set of Requests for Production ("RFPs"), Plaintiffs move for an order compelling production of documents from defendant Oracle.

Plaintiffs' Second Amended Complaint alleges that Oracle uses a number of systems to collect, track, record, profile, and sell Plaintiffs' detailed personal information. Plaintiffs' RFPs—served on November 22, 2022—sought foundational documents necessary to understand the technical "structure, function, and operation" of those systems. *See Exhibit A, Plaintiffs' First Set of RFPs.*<sup>1</sup> After extensive negotiations, the parties agreed on a set of search terms for this production. Yet, while these search terms identified over 285,000 responsive documents, Oracle is withholding the vast majority—over 90%—of those documents from production based on its so-called "relevance review." However, here, a secondary layer of review is unwarranted because the parties' already negotiated the search terms to address Oracle's objections based on, *inter alia*, relevancy and proportionality.<sup>2</sup> As such, the search terms are narrowly tailored to identify responsive, relevant documents in proportion to the needs of the case. Oracle, which refuses to disclose the criteria it employs in making its "relevancy" determinations, can only be seen as engaging in purposeful stonewalling in the production of documents that Plaintiffs plainly have the right to inspect. Accordingly, Plaintiffs respectfully ask that the Court order Oracle to produce all non-privileged documents responsive to the parties' agreed-upon search terms.

For two months, the parties engaged in vigorous negotiations that resulted in a mutually agreed-upon set of search terms. On March 17, 2023, Oracle sent Plaintiffs search terms it intended to use to search Oracle's non-custodial documents as well as corresponding hit counts.<sup>3</sup> Plaintiffs responded with revisions to Oracle's proposed search terms. Following a meet and confer, Oracle came back with a set of revisions to Plaintiffs' search terms and hit counts on April 19, 2023. On April 26, 2023, Plaintiffs requested Oracle explain why it had deleted certain search strings and terms and, for purposes of negotiation, agreed to drop certain search strings. In the interest of resolving the parties' dispute, Plaintiffs also requested that Oracle provide a sample of responsive documents identified by each party's search terms, which Oracle refused. Oracle responded on May 1, 2023, explaining that in its view, the deleted strings were unresponsive, overbroad, and duplicative. The following day, Plaintiffs made further suggestions "tailored to address Oracle's objections" that "address any alleged overbreadth." May 2, 2023 Email from D. Rudolph. On May 9, 2023, Oracle responded, thanking Plaintiffs for "agreeing to the majority of Oracle's revisions to Plaintiffs' search terms," and provided further edits to a particular search string. Again, Plaintiffs responded the next day with edits to that search string. On May 16, 2023, the parties finally agreed to an initial set of significantly narrowed search terms. These final agreed-upon terms yielded hit

---

<sup>1</sup> The search terms did not address RFP No. 2 seeking data related to the Plaintiffs; Plaintiffs have diligently pursued that data through a separate process, as explained in Plaintiffs' Surreply in Opposition to Oracle's Motion to Dismiss (Dkt. No. 99).

<sup>2</sup> Oracle's Responses to Plaintiffs' First Set of RFPs covered the field of typical boilerplate objections about breadth, proportionality, vagueness, relevancy, burden, privilege, and work product. The parties addressed these objections in agreeing upon the search terms.

<sup>3</sup> Oracle later changed course, determining that it would use search terms to collect documents from both custodial and non-custodial sources.

counts, which Oracle shared with Plaintiffs. July 21, 2023 Email from P. Patel (“We are currently reviewing over 285,000 documents that hit on the parties’ agreed-upon search terms.”).

Oracle throws overboard all the work the parties did to agree on search terms and instead employs its own subjective filter on the set of documents generated by the search terms to hand-select—document-by-document—only those documents it is comfortable disclosing. Fourteen months after the requests were served, Oracle has produced only **19,316 documents** out of **over 285,000** that hit on the parties’ agreed-upon search terms, a production rate of just **6.8%**. Oracle overtly intends to keep over 90% of these documents from seeing the light of day, repeatedly conveying the notion that it has “substantially completed production of documents responsive to Plaintiffs’ Set One RFPs.” October 27, 2023 Email from C. Hill.<sup>4</sup> Plaintiffs have repeatedly raised their concerns over Oracle’s unilateral imposition of a secondary review to the agreed-upon search terms. Oracle refuses to explain the parameters of its relevancy review process, stating only that its process was “fully documented” in its objections to Plaintiffs’ RFPs. But those very same objections were already negotiated in reaching a *compromised* set of search terms. Moreover, Oracle’s objections to the RFPs were mostly boilerplate and thus shed no meaningful light on the grounds it now invokes to withhold responsive documents.

Oracle has used this secondary filtration “relevance review” to withhold responsive documents, both across topics and across custodians. Plaintiffs’ RFPs sought technical documents explaining the “structure, function, and operation” of the Oracle Data Marketplace (RFP No. 6) and Data Cloud Core Tag (RFP No. 7) which are at the center of the case and Oracle’s data broker business. Yet Oracle produced in response a scant 307 documents with the term “data marketplace,” and appears to have filtered out many documents that go to the “structure, function, and operation” of the Data Marketplace. Instead, it produced documents that only cursorily mention the data marketplace, such as in a contractual definition in Oracle’s data licensing agreements (58 documents), a one-off sentence in Oracle’s marketing or sales documents (116 documents), documents that refer to *another* company’s data marketplace (15 documents), or are duplicates or near duplicates (28 documents). As for RFP No. 7, Oracle appears to have produced only 74 documents responding to the agreed-upon Core Tag search term string, many of which suffer from the same types of infirmities. By its inadequate production, Oracle has unfairly prejudiced Plaintiffs’ understanding of the “structure, function, and operation” of these two key systems that Oracle used to collect, process, and sell their personal data.

Metrics on Oracle’s custodial production also illustrate that Oracle is withholding responsive documents from many custodians through its unilateral relevance review.<sup>5</sup> In total, Oracle’s production in response to Plaintiffs’ First Set of RFPs includes only 918 emails. The Data Marketplace has been in operation at Oracle since at least 2016, yet only 29 of the 307 documents

---

<sup>4</sup> Oracle has not produced a privilege log, despite months of Plaintiffs’ requests that it do so.

<sup>5</sup> Oracle’s production contains only a nominal number of documents for each of the following custodians despite hundreds or thousands of documents with search term hits for each: Chris J. Bell (Vice President of Product Management Identity for Oracle Data Cloud): 88 documents produced out of 1,934 total custodial documents with search term hits; Chris Tsoufakis (Vice President of Engineering | MOAT (Measurement Products)): 65 out of 5,582 documents; Derek Wise (Chief Product Officer, Oracle Data Cloud): 63 out of 2,581 documents; Chris W. Stark (Vice President of Product Management for Oracle Advertising): 39 out of 2,220 documents; Iain Lowe, (Senior Director of Software Development on the Moat Analytics team): 26 out of 3,890 documents; Benjamin Cassedy (Senior Manager of Identity Graph Core): 23 out of 4,902 documents; Stevan Vlaovic (Senior Vice President of Product and Engineering): 15 out of 965 documents; Mark Dickerson (Vice President of Oracle Advertising): 13 out of 1,528 documents; Matt Carek (Former Senior Director of Privacy & Compliance): 0 out of 407 documents.

containing the term “data marketplace” consist of emails or electronic messages. Despite Oracle’s employees’ prolific usage of the Slack messaging platform, this production contains no Slack messages. This shows that Oracle has applied an inappropriate and overly-restrictive “relevance” criteria. *CrossFit, Inc. v. Nat'l Strength & Conditioning Ass'n*, No. 14-1191, 2019 WL 6527951, at \*15 n.15 (S.D. Cal. Dec. 4, 2019) (“documents [that] hit upon the keywords to which the Parties agreed” are “*presumptively* relevant”). Without relevant custodial documents and emails, Plaintiffs’ ability to identify deponents and prepare for class certification is materially impaired.

During the meet and confers, Oracle’s only defense to the incredibly low percentage of documents produced from the responsive set (about 6.8%) was that a low percentage alone does not demonstrate that its production is deficient. But that position incorrectly accepts the premise that its review was appropriate to begin with, which it was not. Rule 26(b)(1) does not impose an “overly exacting” relevancy standard. *In re Volkswagen “Clean Diesel” Mktg., Sales Pracs., & Prod. Liab. Litig.*, No. MDL 2672, 2018 WL 340640, at \*1 (N.D. Cal. Jan. 9, 2018). Rather, it entitles Plaintiffs to evidence that has “any tendency to make . . . more or less probable . . . [a] fact [that] is of consequence in determining the action.” *Id.* (quoting Fed. R. Evid. 401) (modifications in original). It should be the narrow exception that Plaintiffs do not receive documents that are responsive to the agreed-upon search terms, not the rule an overwhelmingly large percentage (93.2%) of the time. This is especially true when, as here, the parties’ search terms were designed to identify documents that are relevant and proportional to the needs of this litigation. Accordingly, Plaintiffs respectfully submit that Oracle must produce the remainder of the 285,000 documents that hit on the parties’ agreed-upon search terms, excluding only those for which it asserts a privilege. Oracle may not grant itself a second bite at the apple by reapplying its objections and producing only those documents it unilaterally deems relevant. A contrary interpretation would undermine core principles of Rule 26. Plaintiffs therefore request the Court enter an order compelling Oracle to produce all non-privileged documents responsive to Plaintiffs’ First Set of RFPs that hit on the parties’ agreed-upon search terms within fourteen days.

## II. ORACLE’S POSITION

It is axiomatic that a party responding to discovery undertake a relevance review before producing documents, and only produce documents that are responsive to the propounded requests for production. Yet Plaintiffs ask the Court to compel Oracle to produce all non-privileged documents hitting on the parties’ agreed to search terms *regardless* of whether those documents are responsive to Plaintiffs’ RFPs. Plaintiffs contend that any relevance review is improper because the parties negotiated “narrow” search terms. Plaintiffs do not cite any case law to support their position.<sup>6</sup> Nor could they. It is well-established that an “agreement to run a search using the parties’ agreed-upon terms does **not** constitute . . . acquiescence to produce all resulting documents.” *FlowRider Surf, Ltd. v. Pac. Surf Designs, Inc.*, No. 15-cv-1879, 2016 WL 6522807, at \*8 (S.D. Cal. Nov. 3, 2016) (emphasis added). Plaintiffs’ claim that a low responsiveness rate justifies abandoning any relevance review is also contrary to the law. “[A] party’s suspicion that all responsive documents have not been produced, without more, is generally insufficient to warrant an order compelling production.” *Swanson v. ALZA Corp.*, No. 12-cv-04579, 2013 WL 5538908, at \*3 (N.D. Cal. Oct. 7, 2013). Oracle has fully satisfied its discovery obligations with

---

<sup>6</sup> Plaintiffs’ sole case citation in supposed support of a further production concerns a sanctions motion following a documented “mass deletion event” implicating allegedly relevant ESI. (See Plaintiffs’ Position, *supra*, at 3 (citing *CrossFit, Inc. v. Nat'l Strength & Conditioning Ass'n*, No. 14-1191, 2019 WL 6527951, at \*15 n.15 (S.D. Cal. Dec. 4, 2019).) Here, by contrast, Plaintiffs cannot point to a single document or even *category of documents* that Oracle has failed to produce (much less, deleted).

respect to Plaintiffs' RFPs. Plaintiffs' request to deviate from well-established discovery principles should be rejected.

Plaintiffs insist that Oracle's relevance review was "unwarranted" and that they are entitled to all documents that hit on the agreed-upon search terms. That is not the law. Rule 26 permits "discovery regarding any nonprivileged matter that is relevant to any party's claim or defense and proportional to the needs of the case." Fed. R. Civ. P. 26(b)(1). Consistent with this principle, the ESI Order contemplates using search terms "to **narrow** the universe of documents to be searched" and "produce **relevant** information consistent with Fed. R. Civ. P. 26(b)(1)." (ECF No. 48, Section III, VI.2 (emphasis added).) In other words, the Federal Rules and the ESI Order contemplate reviewing documents containing search terms for relevance.

Numerous courts deciding similar disputes agree. *See, e.g., O'Donnell/Salvatori Inc. v. Microsoft Corp.*, 339 F.R.D. 275, 277 (W.D. Wash. 2021) (denying motion to compel production of all non-privileged documents hitting on search terms); *Scherer v. FCA US, LLC*, No. 20-cv-2009, 2021 WL 5494463, at \*2-3 (S.D. Cal. Nov. 23, 2021) (same); *FlowRider*, 2016 WL 6522807, at \*8 (same); *BancPass, Inc. v. Highway Toll Admin., LLC*, No. 14-cv-1062, 2016 WL 4031417, at \*3 (W.D. Tex. July 26, 2016) (same); *Palmer v. Cognizant Tech. Sols. Corp.*, No. 17-cv-6848, 2021 WL 3145982, at \*9 (C.D. Cal. July 9, 2021) (same). An "agreement to run a search using the parties' agreed-upon terms does not constitute . . . acquiescence to produce all resulting documents." *FlowRider*, 2016 WL 6522807, at \*8 (denying "motion to compel . . . all documents that 'hit' on the parties' ESI search terms" because not all results are relevant and responsive).

For example, in *Palmer*, the plaintiff asked the court to compel defendants to produce all non-privileged documents containing at least one search term without conducting a review for responsiveness. 2021 WL 3145982, at \*7. Like Plaintiffs here, the *Palmer* plaintiff complained about a "surprisingly low production rate, considering the parties agreed to only highly relevant search terms." *Id.* The court rejected this argument, acknowledging the broad search terms and false positives led to the low responsiveness rate. *Id.* at \*7-8. Similarly, in *BancPass*, the court did not require the defendant to produce every non-privileged document that contained a search term. 2016 WL 4031417, at \*3. The defendant reviewed 3,489 documents for responsiveness and ultimately produced only 34 of them. *Id.* Even then, the court refused to require the defendant to produce the entire universe of search results, as there was "no reason to believe that [defendant] has withheld documents it was obligated to produce." *Id.*

In fact, not only do courts allow a responsiveness review—they often require it. *See, e.g., Youngevity Int'l Corp. v. Smith*, No. 16-cv-704, 2017 WL 6541106, at \*12 (S.D. Cal. Dec. 21, 2017) (ordering plaintiff to screen for responsiveness and privilege). In *Youngevity*, the court explained that while "[s]earch terms are an important tool parties may use to identify **potentially responsive** documents," they "do not . . . replace a party's requests for production." *Id.* at \*10 (emphasis added). Ultimately, the court concluded that "producing all documents hit upon without reviewing a single document prior to production or engaging in any other quality control measures, does not satisfy [plaintiff's] discovery obligations." *Id.*

Consistent with the Rules, ESI Order, caselaw, and industry practice, Oracle ran the agreed search terms against the agreed data sources, reviewed the results for responsiveness to Plaintiffs' RFPs and for privilege, and produced the non-privileged, responsive documents. (Decl. of Christin Hill ("Hill Decl.") ¶¶ 5-7.) Plaintiffs' claim that Oracle employed a "subjective filter . . . to hand-select . . . only those documents it is comfortable disclosing" is unsupported and false. (Plaintiffs'

Position, *supra*, at 1-2.)

As to the rate of responsiveness, as Oracle explained during the parties' February 13 meet and confer, as well as in a subsequent letter, the relatively low responsiveness rate is a result of the fact that (1) the parties' agreed search terms were overinclusive (like in *Palmer*) and (2) documents concerning the "structure, function, and operation" of the products at issue are not likely to be found in custodial collections. (Hill Decl. ¶ 10.) Indeed, Plaintiffs' RFPs were largely comprised of copy-and-paste verbiage seeking technical overviews as to the function of Oracle's products and services. (*Id.* ¶ 2 (stating that RFP Nos. 4-13 all seek—verbatim—the "format and schema of databases," "pipelines and other data sources used to populate databases," "data anonymization/de-identification procedures," and "data deletion or scrubbing procedures" related to in-scope products).) For each of these requests, Oracle agreed to produce non-privileged documents "sufficient to show the structure, function, and operation" of those products and services, specifically, "the pipelines and other data sources used to populate databases" and any associated "data anonymization/de-identification procedures." (*Id.* ¶ 3.) Once the parties agreed on search terms, Oracle began review of the documents and made responsiveness calls consistent with the scope of documents and information it agreed to produce in its written responses and objections to the RFPs.<sup>7</sup> (*Id.* ¶ 7.) At no point during this months-long process did Plaintiffs raise any objection to Oracle's relevance review. (*Id.* ¶ 6.) When Plaintiffs first raised the subject of Oracle's relevance review—two months after Oracle substantially completed its review—Oracle directed Plaintiffs to its responses and objections to each RFP, which formed the basis of Oracle's review.<sup>8</sup> (*Id.* ¶ 10.) Plaintiffs lack any basis to argue that Oracle acted contrary to its discovery obligations.<sup>9</sup>

Ignoring these facts, Plaintiffs speculate that more responsive documents *should* exist. But Plaintiffs' speculation does not support a reasonable inference that further responsive documents *do* exist, and it does not support compelled production. *See Swanson*, 2013 WL 5538908, at \*3 (mere suspicion that responsive documents are being withheld insufficient to compel production). Rather, Plaintiffs must make "a showing of what specific documents have been withheld." *FormFactor, Inc. v. Micro-Probe, Inc.*, No. 10-cv-03095, 2012 WL 1575093, at \*7 (N.D. Cal. May 3, 2012). They fail to do so.

Plaintiffs present no plausible basis for granting their motion to compel. Their attempt to re-write the rules of discovery should be denied.

---

<sup>7</sup> Plaintiffs misrepresent Oracle's position on search terms. (*See* Plaintiffs' Position, *supra*, n.3.) Oracle never agreed to use search terms only for non-custodial sources. Rather, it agreed to investigate whether any documents responsive to Plaintiffs' RFPs could be located through a targeted search. After determining that it was not possible to do so, Oracle indicated that it would apply search terms across all sources. (Hill Decl. ¶ 4.)

<sup>8</sup> Contrary to Plaintiffs' argument, Oracle did not "refuse[] to disclose the criteria it employs in making its 'relevancy' determinations." (Plaintiffs' Position, *supra*, at 1.) Oracle is not obligated to disclose additional, privileged information on the instruction it provided to its reviewers.

<sup>9</sup> Moreover, the information Plaintiffs seek on the "structure, function, and operation" of relevant products and capabilities is more readily discoverable (and understandable) through a Rule 30(b)(6) deposition. *See Rodriguez Ayala v. Cnty. of Riverside*, No. 16-cv-686, 2017 WL 2974919, at \*3 (C.D. Cal. July 12, 2017) (denying production in response to overly broad discovery requests where information could be obtained through deposition testimony). Despite that, Plaintiffs have steadfastly refused to seek information through these means.